Artificial Intelligence for Gibbsite Crystallization Control at CBA

Thiago Franco[1], André Arantes[2], Anderson Duck[3], Michel Santos[4], Hugo Teixeira[5] and Kenji Urazaki[6]
1. Process Engineer
2. Process Engineer
3. Process Consultant
4. Data Architect
Companhia Brasileira de Alumínio, Alumínio, Brazil
5. Data Consultant
6. Data Analyst
Votorantim SA, São Paulo, Brazil
Corresponding author: thiago.franco@cba.com.br

Abstract

Industry 4.0 is a new era in the context of major industrial revolutions. It encompasses the main technological innovations in the fields of automation, control and information technology, simulation, big data and Internet of Things applied to manufacturing processes. Several changes are taking place in the industrial processes, having impacts in all market sectors. The impacts of Industry 4.0 on productivity, cost reduction, employee safety, control and optimization over the production process, among others, boost company's competitiveness by improving machine efficiency, reducing maintenance costs, reducing energy and raw material consumption, as well as improving product quality.

In this context, one of CBA's focuses is the development of artificial intelligence models. The aim is the reduction in process variability and the improvement on the produced alumina granulometry. An AI model was created to predict the hydrate precipitation process behavior. The paper presents the implementation of the AI model on the CBA precipitation process, the utilization of this tool in the daily routine of process engineering and the main results on process control and alumina quality.

Keywords: Artificial Intelligence, Industry 4.0, Gibbsite Crystallization, Hydrate Precipitation.

1. Introduction

Gibbsite crystallization is a process widely studied in universities and research centers around the world, reflected in the great number of patents since Bayer process beginnings. Today, it exists in more than a hundred refineries with different production capacities that vary from 0.1 to 6.4 million tonne per year. The residence time of the crystallization process ranges from 30-100 hours, where high yields are required to justify economic viability. To produce standard quality alumina with high plant yields is a challenge to every refinery, according to Stamatiou et al[1]. In this context, process modelling with computational resources are important tools for process control.

1.1 Gibbsite Crystallization in the CBA's Alumina Refinery

Pregnant liquor (PGL) from the Heat Interchange Departments (HIDs - stage responsible for cooling the red side liquor) flows to the agglomeration tanks where fine aluminum hydroxide seed is added. After this stage, the liquor and agglomerated solids flow to the growth stage, controlled to target temperature by cooling systems, called ISCs (Interchange Stage Cooling) composed of

flash tanks, barometric columns and cooling towers. The growth tanks also receive coarse seed from the classification system. The resulting suspension is pumped to the last growth tanks in the chain, without seed addition and later to the hydrocyclone classification system.

The first hydrocyclone clusters separate the final product from the seed which, in turn, are separated between fine and coarse seed by new hydrocyclone batteries. Both product and seed are filtered to recover liquor, while the product is then sent to calciners, the fine seed to the agglomeration tanks and the coarse seed to the growth tanks. The precipitation and classification system at CBA are illustrated in figure 1.

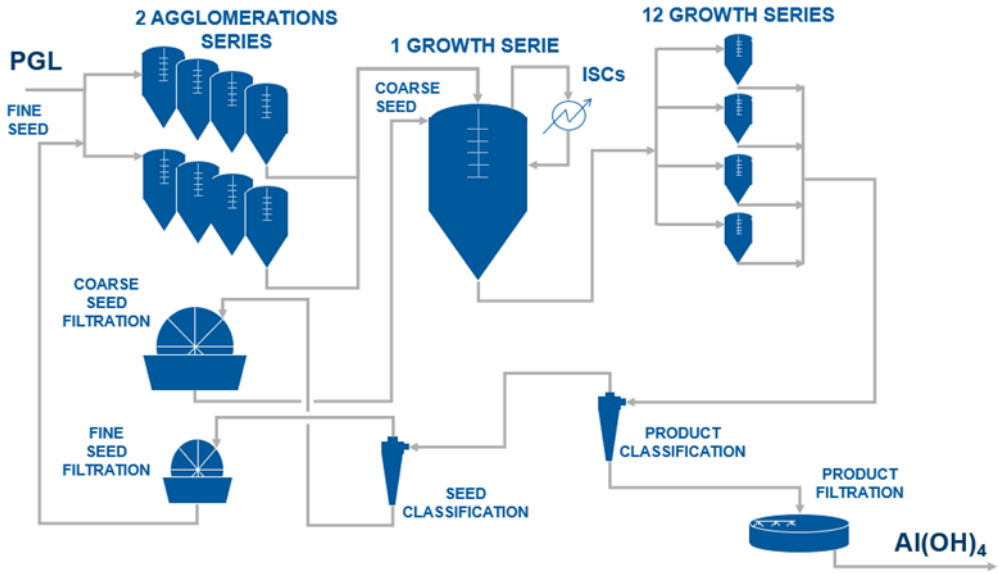

Figure 1. Flowsheet of the CBA precipitation process.

There are a total of 75 tanks. Ten (10) of these tanks are agglomeration tanks divided into 2 parallels series. The agglomeration tanks feed a series of 7 growth tanks connected to ISCs and 12 parallels series of 58 growth tanks. The total operation volume varies between 55 000 and 65 000 m³. In these tanks, aluminum hydroxide is crystallized through the chemical reaction as follows:

$$Al\,(OH)_4^- + 2Na^+ \xrightarrow[+\,Seeds]{Cooling} Al\,(OH)_3 + 2\,NaOH$$

This is the main reaction responsible for plant yield, alumina strength and quality indicators. The reaction occurs through three mechanisms: nucleation, growth and agglomeration.

There are different kinetics for the three mechanisms. The aluminum hydroxide solubility in the liquor (dependent on the dissolved chemical species) and all the process parameters to be controlled including temperature and solid concentration profiles make the crystallization process very complex.

1.2 Artificial Intelligence (AI)

To model the chemical process is a challenge faced by every hydrometallurgical plant, and in the context of industry 4.0, new applications are emerging. The artificial intelligence models problems through nonparametric algorithms using a method known as machine learning. It

generates general structures and hyperparameters which are constantly adjusted based on process data.

There are primarily two ways of modelling a chemical process [2]:

a. Parametric: the model uses a set of equations (linear, nonlinear, differential, and algebraic) which are derived from physical laws, for example, mass conservation and empirical knowledge from the process. The model structure is well defined and depends on the process parameter definition by using a limited amount of data from the studied process.
b. Nonparametric: arbitrary models that are more flexible and normally depends on a vast amount of data from the process. In this approach, previous knowledge of the studied process is not needed, making the models development and input/output data system simpler, resulting in faster computational run times.

There are different methods of developing machine learning [3]:

a. Supervised training: the inputs and outputs are known for all training data sets. The training algorithm works by predicting the process behavior for each process output and comparing it to real data. The difference between them is used to update the hyperparameters of the model and the training algorithm runs until the difference is acceptable.
b. Unsupervised training: there is no need for outputs. This type of learning is used to segregate data by featuring similarities. The number of clusters/classes can be chosen in the algorithms.

Many types of machine learning models exist, for example decision trees, linear regression, neural networks and random forests, each one with advantages and disadvantages. They utilize comparison methods (like least square method), and the training of the machine to find the best hyperparameter fit is based on the trial-and-error method. In the development of an artificial intelligence model, it is important to have a data specialist to support with modelling methods knowledge and a technician to provide information and insights about the chemical process. Together, they can synergically build appropriate solutions to provide a predictive model. In this way, Votorantim SA and CBA integrated their teams to develop an AI model to predict the behavior of the gibbsite crystallization process.

In this work, the AI model was developed using Neural Networks (NNs). NNs uses powerful machine learning to capture the nonlinear data correlation or complex patterns in the process, They perform better than classical models by increasing nonlinearity and the number of process variables under consideration to resolve the problem [4]. They are usually used in many different types of problems, for example, to fit the model to experimental data, to make machine diagnoses, for pattern recognition and product quality control, among others. They also have a robust response even in the presence of disturbances while providing fast run times [2]. The neural network used in this work belongs to the recurrent class, that means the signal of an input always uses a prior feedback.

The operation structure uses interconnected neurons (or nodes), in which each connection between nodes has a weighted value and adjusted (the hyperparameters) with the new inputs. Internally, nodes have functions called transference functions composed by the hyperparameters, which can be any mathematical function [2]. This structure is organized in layers: the first layer (or input layer) is the set of neurons that receive the input data; the other layers (one or many others in sequence) called hidden layers is the set of neurons that receive the outputs from the input layer; and the output layers which aggregates the output from the last hidden layer and generate the final output of the network.

The hyperparameters of the neural network in the transference functions cannot be calculated directly by an analytical formula. They are set by an interactive method of training, using the training data set of the process, composed by the inputs and outputs, and adjusting the created algorithm [2].

The LSTM (Long Short-Term Memory) neural networks were used for this problem due to storage and information access [5]. The neuron of a LSTM is called memory cells and they are structured in three decision gates to maintain the information flux [5]. It is a challenge to train LSTM: the input data set needs to be normalized and stored to a specific standard, the removal of outliers, having the training data set at a specific size, among others for the modelling success.

1.3 Control Philosophy

For alumina quality control and crystallization yield it is necessary to make daily adjustments to regulate agglomeration, growth and nucleation. There are several variables that regulate these mechanisms, such as temperature, solids concentration, caustic concentration of liquor, liquor flow, residence time, etc. Finding a stable profile of these variables results in a stable particle size distribution at the end of the process.

These variables must be managed within the control limits, such as temperature profile and solids concentration, to control the nucleation and agglomeration rates within a target range. ISCs play an important role in temperature control in growth tanks. Temperature regulates the supersaturation of the liquor, generating crystallization yield and promoting the necessary nucleation. The fines generated in the growth tanks are consumed in the agglomeration stage. Results outside control limits can bring negative impacts on alumina quality. These include the generation of excess fine material pushing the product quality outside of specifications.

We adopted a model based on the particles in the fine seed of size fraction -25μm/-44μm sieves. This sample is an indicator of the change in the fine particle population of the circuit and if the process parameters are not adequately controlled.

2. Implementation of Artificial Intelligence for Gibbsite Crystallization Control

An AI model was created with the aiming of reducing process variability and improving the produced alumina granulometry. The model was created to predict the crystallization process behavior. The AI application processes more than 260 process variables and considers

1. Granulometry control cycle response time of 5 to 10 days.
2. Assertively bringing the prediction of the fine seed granulometry profile.
3. Providing safety in decision making on the main control variables.
4. Reducing variability in the granulometry of alumina.
5. Ensuring the quality index provided to clients.
6. Providing conditions to exploit yield in the entire crystallization circuit.

2.1 Method and Data Architecture

Microsoft Azure resources were used, and the model was written with Python programming. The hydrate crystallization from CBA was modeled through NNs and generates a 7-day prediction using the daily average of process variables of the prior day. All the execution steps are controlled by Data Factory, a plugin from Azure. The NNs modelling used Keras Library from Python, which is very user-friendly and has many functions for training and tuning the neural network. The Root Mean Squared Error method (RMSE) was used to tune the model outputs where the

first and second derivatives were evaluated. The derivatives were calculated using the Savitzy-Golay filter. Figure 2 presents a graph of this learning process of the AI model.

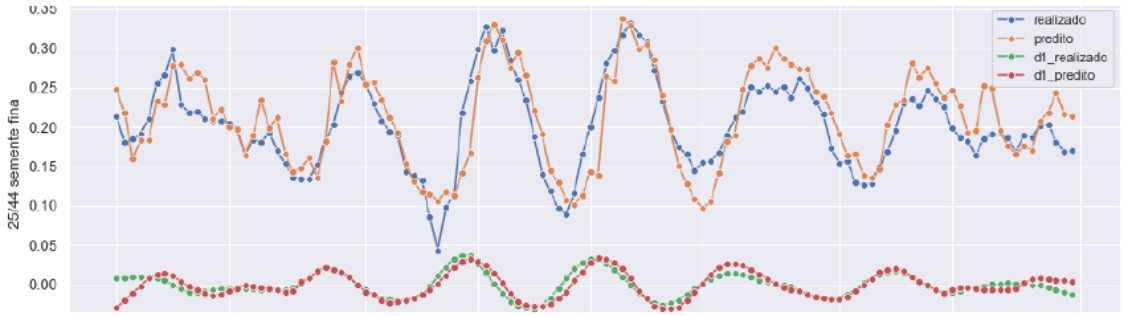

Figure 2. Minus 25/44 µm relationship of fine seeds from real values and AI model predictions and first derivatives in the learning process

In the current solution, the process data is extracted from an Excel spreadsheet on a shared server, which uses the PI Datalink extension to extract the values from the historized process data server (PIMS), where the TAGs are sequenced and queued for use based on the frequency of capture by the sensors. The data acquisition uses the ETL resources combined between the Data Factory and the Event Hub, forming the first stage where the data presents a raw and semi-structured format.

In the next step, the information is stored in Data Lake, producing an accurate version, enriched with other complementary information from other PowerApps solutions. Then, data is transferred to a SQL server where it is easily accessible for the Python algorithm. The data is categorized and then an embedding method is applied in the Machine Learning stage. Missing data is treated as a '-1' input and all variables are normalized using a suitable method from Sklearn Library, another plugin from Azure.

The modified process data is fed daily to LSTM together with data from days before. The number of days inputted to the NNs was chosen with the process engineers guidance incorporating the process dynamics and knowledge. In this way, process indicators are predicted in a range of seven (7) days after the last days average. This horizon was set up to offer precise information about the tendency in the crystal granulometry profile, to improve the quality and to assertiveness control the process.

2.2 Daily Utilization

Every day, after the crystallization process parameters are analyzed and historized in an excel spreadsheet, the methodology described above is applied and an automatic report is generated in Power BI application. Predictions of the fine seed relation -25µm/-44µm sieves from 1 to 7 days ahead are given and carefully monitored by process engineers. Figure 3 shows the comparison between the indicator values (in orange) and the predictions (in blue).

When the predictions of the relation -25µm/-44µm in the fine seeds shows a tendency to increase, it indicates that the fine population will increase in the beginning of the process. So, it is important to set the crystallizers to consume fine population or to produce fewer fine particles when this consumption power is weak. On the other hand, if the predictions of the relation -25µm/-44µm shows a tendency to decrease, the particles are getting coarser. The actions taken must be to avoid decreasing the specific surface area in the particles, possibly causing superfines nucleation later.

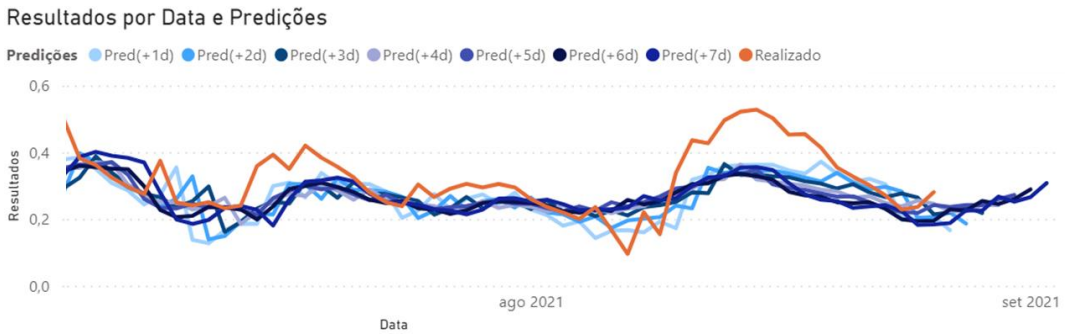

Figure 3. Minus 25/44 µm relation of fine seeds for real values and AI model predictions in daily routine

Along with the graphs, a table was also developed to show the main process parameter disturbances to explain the behaviors of the prediction curves, as well as to highlight the day these disturbances occurred.

The model constantly improves as it learns from new process events or with deviations between the model and reality.

3. Results

CBA's Alumina Refinery is already enjoying the AI model benefits. It was first introduced to the process engineering team on 1st July of 2020. There was a benefit that was observed for the environmental. By producing alumina with lower granulometry, less residue was generated in the calciners and in the gas treatment of the smelter reduction process. Figures 4 and 5 demonstrates the difference of the alumina attrition index and alumina passing granulometry on 44 µm sieve one year before and one year after implementation

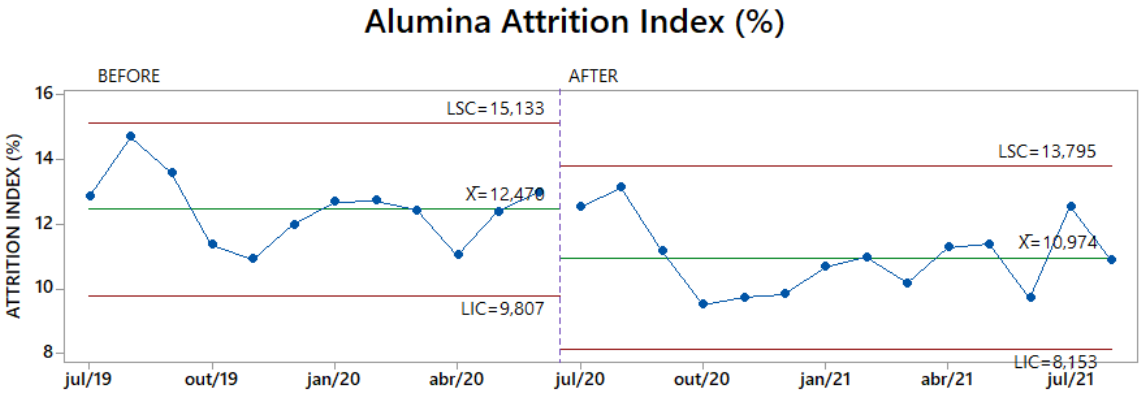

Figure 4. Control chart for the alumina normalized attrition index improvement after artificial intelligence application

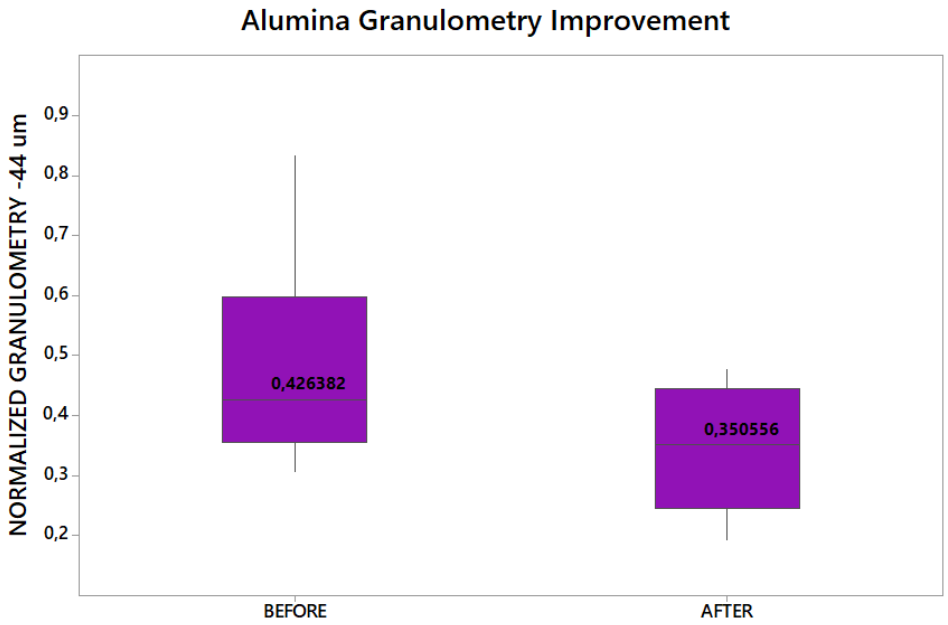

Figure 5. Boxplot of alumina normalized granulometry improvement with artificial intelligence application

Alumina attrition index showed an improvement of 12%, indicating a better efficiency in the agglomeration control in crystallization, while the alumina granulometry in the -44 μm sieve showed improvement of 30% and more than 10% in process variability, indicating a much better control of the entire crystallization process.

4. References

1. Evangelo Stamatiou, David Chinloy, Bekir Çelikel, Murad Kayaci, and Esra Savkilioglu, Hatch – ETI Aluminyum Precipitation Modeling, *Light Metals* (2013), p. 143-146
2. David M. Himmelblau, Accounts of experiences in the application of artificial neural networks in chemical Engineering. *Industrial & Engineering Chemistry Research*, 2008, v. 47, n. 16, p. 5782 – 5796.
3. Ivan Nunes Silva et al., Artificial Neural Networks: A Practical Course, Springer, 2017.
4. Imad Basheer, M. Hajmeer, Artificial neural networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods*, 2000, v. 43, n. 1, p. 3-31.
5. Luisa Castello Branco de Sá, Desenvolvimento de um modelo de predição de temperatura e umidade de uma estufa baseado em aprendizado de máquina [dissertation]. *Belo Horizonte: Universidade Federal de Minas Gerais*, 2020.